

УДК 541.69

## ВОЗМОЖНОСТИ И ПЕРСПЕКТИВЫ КОНСТРУИРОВАНИЯ БИОЛОГИЧЕСКИ АКТИВНЫХ ВЕЩЕСТВ

*Раевский О. А., Сапегин А. М.*

Проанализировано современное состояние исследований по проблеме связи структуры веществ с проявляемыми ими свойствами, главным образом — биологической активностью, установление которой представляется единственной разумной альтернативой дорогостоящему тотальному скринингу. Показано, что в этой проблеме, затрагивающей обширные разделы химии, физики, математики, биологии, дальнейший прогресс возможен только при гармоническом развитии и совершенствовании языка описания структуры в рамках использования ЭВМ, методов выявления связи структура — свойство, количественных критериев свойств.

Библиография — 171 ссылка.

### Оглавление

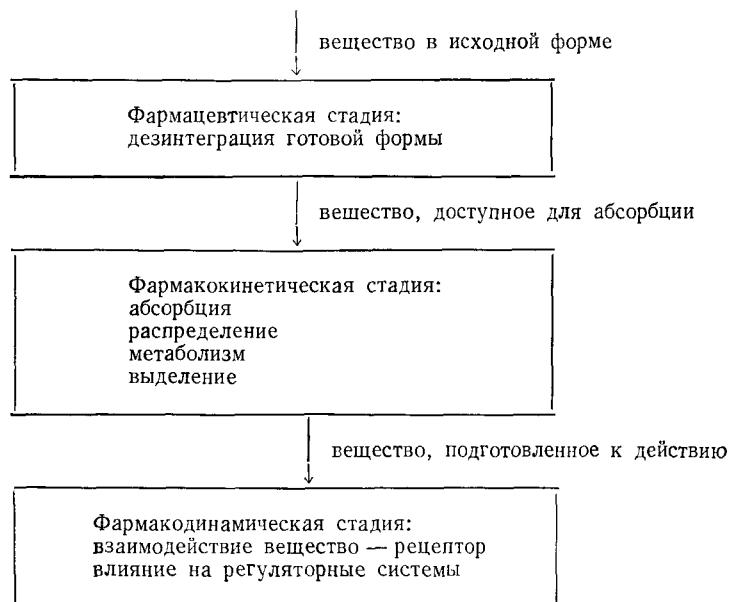
|   |      |
|---|------|
| I. Введение . . . . .   | 1565 |
| II. Молекулярное распознавание — платформа связи структура — активность . . . . .                                     | 1567 |
| III. Корректное описание структуры в рамках использования ЭВМ для установления связи структура — активность . . . . . | 1569 |
| IV. Методы выявления связи структура — активность . . . . .   | 1573 |
| V. Комплексы программ по выявлению связи структура — активность . . . . .   | 1578 |
| VI. Заключение . . . . .  | 1581 |

### I. ВВЕДЕНИЕ

Научно-технический прогресс сопровождается экспоненциальным ростом числа химических соединений, применяемых в промышленности, сельском хозяйстве, медицине, быту. Химия, благодаря практически беспределным возможностям синтеза новых соединений, обеспечивает постоянное улучшение полезных свойств используемых веществ. В настоящее время число известных химических соединений оценивается примерно в десять миллионов. Однако все они составляют только ничтожно малую часть бесконечно большого числа возможных соединений. Об этом, в частности, свидетельствуют и данные, приведенные в [1]: если использовать хотя бы пятьдесят хорошо охарактеризованных в органической химии заместителей для подстановки в различном сочетании в семь свободных положений только одной нафтойной кислоты, можно получить  $7,8 \cdot 10^{12}$  химических соединений. Отсюда явствует высокая вероятность синтеза соединений с гораздо лучшими, по сравнению с существующими, свойствами. В связи с этим встает очень важная проблема, как подойти к решению задачи получения высокоэффективных препаратов в кратчайший срок и с минимальными затратами. Очевидно, что использование в качестве основы для поиска новых препаратов всесторонних испытаний свойств, т. е. тотального скрининга, приведет к немыслимым затратам (при реализации такого подхода для внедрения одного вещества необходим синтез и всесторонние испытания 12—15 тыс. химических соединений). Все изложенное выше ставит химическую науку перед необходимостью выработки чистых принципов конструирования веществ с заданными свойствами.

Согласно [2, 3], формирование представлений о структуре вещества было тесно связано с выявлением взаимосвязи между его строением и проявляемыми свойствами: вначале — на качественном уровне [3, 4], а с открытием периодического закона элементов происходит постепенный переход к установлению количественных соотношений структура — свойство [5].

Особый интерес и значимость проблема установления связи структура — свойство приобретает по отношению к поиску биологически активных веществ (БАВ). Традиционные подходы, такие как выделение БАВ из природных источников и синтез их аналогов, изучение средств народной медицины, массовые биологические испытания и т. д., не в состоянии обеспечить быстро растущие потребности в новых препаратах. Поэтому с середины XIX века начинает развиваться новое направление в поиске БАВ — выявление связи между структурой вещества и проявляемой им биологической активностью (ССА). Истоки этого направления связаны с работами Ричардсона, которому впервые удалось на качественном уровне связать структуру исследуемых органических соединений с проявляемой ими физиологической активностью [6], а также, Крум-Брауна и Фрезера, предпринявшими первую попытку создания математической модели, описывающей связь структура — активность [7]. К настоящему времени (историю развития ССА см. [3]) сформулированы основные принципы установления количественных соотношений структура — активность (КССА). Схематически современные представления о механизме действия БАВ, положенные в основу КССА, можно изобразить в следующем виде [8]:



Только на фармакодинамической стадии взаимосвязь структуры вещества и биологического эффекта, вызываемого его введением в организм, проявляется в полной мере и является специфичной. Что же касается других процессов, связанных с трансформацией веществ, ингибированием ферментов, диссипацией веществ на различных органах и т. д., сведения о них (хотя и постоянно расширяются [9]) носят подчас весьма ограниченный характер, затрудняя, тем самым, выявление КССА. Таким образом, проблема КССА является частным случаем проблемы связи структура — свойство. Однако сложности приложения ССС к биологическим объектам заставляют выходить за традиционные рамки проблемы структура — свойство и уделять внимание не только методам выявления связи, совершенствованию экспериментальных исследований структуры и форм ее описания, но и вопросам количественного определения величины биологической активности, детализации описания самих биологических процессов.

В настоящее время число публикаций, посвященных анализу КССА, исчисляется уже тысячами. В специальном журнале QSAR (Quantita-

tive Structure-Activity Relationships), издаваемом с 1981 г., наряду с оригинальными статьями, публикуются многочисленные рефераты работ по КССА, опубликованных в других изданиях, в том числе монографий и обзоров [10—20]. Однако в подавляющем большинстве такие работы носят либо специализированный характер и посвящены какому-либо одному подходу или вопросу — регрессионному анализу [18], использованию ЭВМ [19], приложению КССА к анализу пестицидной активности [20], либо основное внимание уделяется описанию математических методов анализа КССА [16, 17].

В настоящем обзоре акцент сделан на основополагающие принципы конструирования веществ с заданными свойствами [21, 22], составляющими физико-химический фундамент проблемы установления количественных соотношений структура — активность. При этом, предполагается, что исследователь располагает надежной количественной информацией о величине биоактивности соединений [18]. В этом случае, задача установления КССА сводится к выбору модели, описывающей взаимодействие вещества с рецептором, и метода выявления связи между корректным (в рамках выбранного метода) представлением структуры соединения и проявляемой им активностью.

## II. МОЛЕКУЛЯРНОЕ РАСПОЗНАВАНИЕ — ПЛАТФОРМА СВЯЗИ СТРУКТУРА — АКТИВНОСТЬ

В настоящее время стало очевидным, что при анализе самых разнообразных свойств химических соединений необходимо рассматривать их межмолекулярные взаимодействия с возможными партнерами. Многочисленные факты, свидетельствующие о том, что определенная молекула  $A_0$  (распознаватель) реагирует или образует комплекс в растворе только с определенным молекулярным эффектором  $B_0$  из смеси иногда очень большого числа возможных эффекторов  $B_i$ , легли в основу так называемой «проблемы распознавания в химии» [23—25].

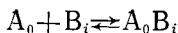
В хронологическом порядке с качественной стороной проблемы распознавания первыми столкнулись химики-аналитики при поиске веществ, эффективно связывающих определенные ионы в многокомпонентных растворах. С развитием химической науки проблема распознавания распространяется на органическую и физическую химию. Здесь уместно упомянуть о фундаментальных работах Ван-Гоффа по созданию основ теории растворов [26], что можно рассматривать как переход к изучению количественной стороны проблемы распознавания. Как отмечено в [25], при количественном описании процессов точность распознавания химических соединений увеличивается на порядок в последовательности от избирательного донорно-акцепторного взаимодействия или Н-связывания молекул, содержащих по одному электронодонорному и акцепторному центру, и, далее, к избирательному комплексообразованию ионов металлов с полидентатными лигандами, избирательным катализитическим реакциям, вплоть до ферментативных реакций. В этой же последовательности увеличивается размер и сложность структуры молекул, участвующих в процессе распознавания.

Методологические принципы проблемы распознавания в химии преимущественно изучаются в настоящее время на примерах избирательного комплексообразования микроциклических полиэфиров и криптандов (см., например, [27, 28]). Однако даже для этих не очень сложных соединений стабильность их комплексов зависит от многих факторов, что затрудняет понимание принципов молекулярного распознавания. Поэтому необходимо четко сформулировать основные положения процесса распознавания.

Согласно [23—25], распознаватель  $A_0$  должен связываться с молекулами определенного типа  $B_i$ , т. е. правильным эффектором из раствора, содержащего  $N$  типов молекул:  $B_0, B_1, B_2, \dots, B_{N-1}$  произвольной концентрации. Обозначим через  $P_f$  количество возможных фальшивых рас-

познаваний. Тогда степень распознавания может быть охарактеризована числом правильных распознаваний, зависящим от общего числа возможных эффектов  $N$  и доли фальшивого распознавания  $P_f$ .

В простейшем случае образования комплекса  $A_0B_i$



с константой реакции

$$\frac{[A_0B_i]}{[A_0][B_i]} = K_{0i} = \exp(-\Delta G_{0i}/RT); i = 0, 1, \dots, N - 1$$

число возможных фальшивых распознаваний  $P_f$  будет равно:

$$P_f = \frac{\sum_{i=1}^{N-1} [A_0B_i]}{\sum_{i=1}^{N-1} [A_0B_0] + \sum_{i=1}^{N-1} [A_0B_i]} = \frac{\sum_{i=1}^{N-1} [B_i] K_{0i}}{[B_0] K_{00} + \sum_{i=1}^{N-1} [B_i] K_{0i}},$$

из чего следует, что распознавание будет тем точнее, чем больше величина отношения констант связывания правильного распознавания к фальшивому или, иначе, чем больше разность свободных энергий ( $\Delta\Delta G$ ) образования комплексов  $A_0B_i$  ( $\Delta G_{0i}$ ) и  $A_0B_0$  ( $\Delta G_{00}$ ) соответственно:

$$\Delta\Delta G_i = \Delta G_{0i} - \Delta G_{00}.$$

Таким образом, высокоспецифичное распознавание требует довольно большого различия  $\Delta\Delta G$  между способностью распознавателя  $A_0$  взаимодействовать с правильным эффектором  $B_0$  и с фальшивыми. При достаточно высоких значениях  $\Delta\Delta G$  распознаватель  $A_0$  способен различать один эффектор  $B_0$  из  $N$  других эффекторов  $B_i$ . Тем самым, проблема целенаправленного поиска веществ с заданными свойствами может быть сведена к конструированию молекул, строение которых обеспечивает максимальное связывание с распознавателем, ответственным за рассматриваемое свойство.

Эти положения проблемы распознавания по отношению к биологическим объектам получили дальнейшее развитие в работах Фишера, который еще в 1894 г. установил, что на характер взаимодействия субстратов с ферментом существенное влияние оказывает их пространственная совместимость [29]. Это легло в основу так называемого принципа «ключ—замок» или «принципа комплементарности». Указанный принцип в настоящее время трансформировался в общепринятое положение, что образование фермент-субстратных комплексов, играющих важную роль в ферментативном катализе, определяется тонкой структурой активного и эффекторного центров и уникальной структурой всей молекулы фермента, обеспечивающими высокую катализическую активность и специфичность действия биокатализаторов [30].

Другой важный вклад в развитие направления структура — активность связан с именем Хэнча, который в шестидесятых годах нашего столетия предложил эмпирическую модель биологической активности, основанную на линейной зависимости свободной энергии рассматриваемого процесса от физико-химических параметров соединений. Эта модель является наиболее распространенным инструментом в работе по оптимизации биологического действия химических соединений [1, 31].

До последнего времени развитие идей Фишера и Хэнча сформировало в задаче структура — активность два независимых подхода, связанных с отбором общих элементов структуры в различных соединениях с одним видом активности и с оптимизацией величины биоактивности в гомологических рядах химических соединений. В дальнейшем будет показано, что переход на качественно новый уровень понимания проблемы КССА требует объединения этих подходов в рамках единого системного исследования. Такая постановка задачи подразумевает доста-

точно полное и корректное описание структуры соединений, включая особенности электронно-топографического и конформационного строения молекул.

### **III. КОРРЕКТНОЕ ОПИСАНИЕ СТРУКТУРЫ В РАМКАХ ИСПОЛЬЗОВАНИЯ ЭВМ ДЛЯ УСТАНОВЛЕНИЯ СВЯЗИ СТРУКТУРА — АКТИВНОСТЬ**

Прежде чем приступить к исследованию связи между структурой и активностью необходимо выбрать способ представления исходных данных [14]. Было затрачено много усилий на подбор доступной экспериментально или путем расчетов системы параметров, которая достаточно полно описывала бы структуру. В настоящем обзоре рассматриваются четыре основных подхода к описанию структуры соединений.

#### **1. Фрагментарный код суперпозиции подструктур**

Фрагментарный код суперпозиции подструктур (ФКСП) был предложен в начале семидесятых годов [32]. При анализе химической структуры на основании этого подхода выделяются потенциально активные центры, названные «дескрипторными центрами» (ДЦ), к которым отнесены атомы, содержащие подвижные  $p$ - и  $\pi$ -электроны или целый электростатический заряд. Это все гетероатомы: азот, кислород, сера, фосфор, галоиды, металлы, а также, циклы и пары атомов углеродов, связанные кратными связями. Подробный список ДЦ приведен в [33]. Первоначально для каждого соединения на основании его структурной формулы составлялся набор линейных дескрипторов, каждый из которых представлял собой семизначное число (по две цифры отводились на код дескрипторного центра и длину цепи; признак сопряжения выражался одной цифрой — равной единице при наличии сопряжения и нулю — в противном случае). Учет специфики циклических фрагментов при кодировании соединений для анализа их сходства проведен в [34]. Свое применение ФКСП нашел в статистических исследованиях информационного массива биологически активных соединений. В дальнейшем предложен способ кодирования структуры химического соединения, использующий понятие ДЦ, но сохраняющий топологию молекуллярной структуры как единого целого в форме матрицы связности дескрипторных центров (МСДЦ) [33]. В этой матрице диагональными элементами являются коды ДЦ, а недиагональными элементами могут быть значения расстояний между соответствующими ДЦ (в случае матрицы геометрии ДЦ) или число связей между ними. Этот подход использован в ряде работ [16, 17] для выявления общих элементов структуры соединений, обладающих конкретным видом биологической активности.

#### **2. Системы топологических и геометрических дескрипторов**

Естественным развитием языка описания химических структур является появление систем топологических и геометрических дескрипторов. Среди линейных способов кодировки структур наибольшее распространение получила система Виссессера [35]. Однако в силу ограничений, присущих линейным системам кодирования, о главном из которых уже упоминалось, в последнее время предпочтение отдается кодированию структур с помощью матриц связности с широким набором молекуллярных дескрипторов. Это прежде всего, топологические дескрипторы, включающие в себя: дескрипторы фрагментов, подструктурные дескрипторы, дескрипторы связности, окружения и т. д. Все они предназначены для более точного описания окружения активных центров молекул [14]. Недостаток этого способа кодировки, также как и кода ФКСП, — в том, что кодируется не структура химических соединений, а их структурная формула. Некоторого улучшения этого способа описания структуры можно добиться, если наряду с топологическими дескрипторами исполь-

зователь и геометрические. Последние, описывая общую форму и размер молекул, создают представление о трехмерной структурной диаграмме соединения. Более полному описанию структуры соединений способствуют конформационные расчеты равновесных расстояний между активными центрами методами молекулярной механики [36]. Однако и в этом случае, остаются невыясненными электронные свойства активных центров, определяющие их способность к межмолекулярному взаимодействию. Это в существенной мере сказывается на выявлении истинных взаимосвязей между структурой химических соединений и проявляемой ими активностью.

### 3. Квантовохимические методы, используемые для описания структуры

Квантовохимические расчеты структуры молекул используются в проблеме структура — свойство с момента их возникновения и в настоящее время, благодаря прогрессу в области вычислительной техники, получают все большее распространение в исследованиях по установлению связи между структурными параметрами молекул и их свойствами, особенно биологическими [19, 21, 37—39].

Существующие к настоящему моменту программы для ЭВМ охватывают практически весь спектр квантовохимических расчетов, начиная от простых схем полного пренебрежения дифференциальным перекрыванием (CNDO) до строгих непараметрических подходов: CNDO/2, CNDO/BV, CNDO/S, INDO, INDO(S), MINDO/2, MINDO/3, MNDO, VRDDO, *ab initio* STO-39. Однако привлечение квантовохимических расчетов к проблеме ССА сводится, в основном, к определению наиболее устойчивых конформаций молекул, в которых активные центры доступны для взаимодействия с рецептором, или к оценке молекулярных физико-химических характеристик, связанных с транспортом вещества до биомассы. Перенесение центра «тяжести» квантовохимических расчетов на электронные свойства молекул позволило бы ставить задачу выявления (конструирования) веществ с заданными свойствами среди принципиально новых классов соединений. Такими возможностями не обладают подходы, описанные выше.

Перспективность использования квантовохимических расчетов для проблемы ССА очевидна. И дальнейшее совершенствование вычислительной техники, переход на неэмпирические расчеты [40] безусловно приведут к возрастанию роли квантовой химии в установлении взаимосвязей структура — свойство. В настоящее же время расчет электронного строения больших рядов химических соединений даже при использовании полуэмпирических методов трудно реализуем из-за огромных затрат машинного времени. В связи с этим, возникает задача создания такой расчетной схемы, которая была бы достаточно проста и, вместе с тем, пригодна для корректного описания электронной структуры молекулы в рамках проблемы ССА. В [41] предпринята попытка наполнить матрицу связности дескрипторных центров квантовохимическим содержанием, учитывающим электронное строение молекул.

Следует, однако, отметить, что квантовая химия далеко не единственный источник информации об электронно-топографическом строении молекул. Более того, в силу перечисленных трудностей, получаемая в результате квантовохимических расчетов информация не всегда достаточно надежна и точна. Поэтому при дальнейшем развитии работ по выявлению взаимосвязей структура — свойство, наряду с этими расчетами, будут совершенствоваться физико-химические подходы к проблеме ССА, прямым образом связанные с экспериментальными исследованиями физико-химических свойств молекул или их фрагментов.

#### 4. Физико-химический подход к проблеме связи структура — активность

Начало широкого применения физико-химических характеристик в проблеме ССА связано с работами Хэнча с соавт. [1, 5, 31]. Описан целый набор разнообразных параметров, используемых при расчетах по уравнению Хэнча [18, 42]. Все эти параметры могут быть разделены на три группы: в первую можно включить параметры, связанные преимущественно с гидрофобными свойствами — логарифм коэффициента распределения ( $\lg P$ ), параметр липофильности ( $\pi$ ), параметр тонкослойной хроматографии ( $\lg K'$ ), парабор ( $Pr$ ), мольный объем ( $V_m$ ), молекулярный вес ( $M$ ), логарифм распределения в воде ( $\lg C_{aq}$ ), парциальное давление ( $p$ ), параметр растворимости Гильдебранда ( $\sigma_s^*$ ), мольная аттракционная постоянная ( $F_a$ ), энтропийная константа ( $\sigma_s^{**}$ ) и др. Во вторую группу входят различные индукционные и резонансные константы ( $\sigma^*$ ,  $\sigma^0$ ,  $\sigma^I$ ,  $\sigma^K$  и т. д.), радикальные константы заместителей ( $E_R$ ), параметры Свайна и Лаптона ( $F$ ,  $R$ ), поляризумость ( $\alpha$ ), молекулярная рефракция ( $MR$ ), дипольный момент ( $\mu$ ), потенциал ионизации ( $IP$ ), частоты в спектрах ( $v$ ), константы ионизации ( $pK_a$ ,  $pK_b$ ) и другие характеристики, связанные с электронными свойствами. Наконец, третью группу составляют параметры, характеризующие стерические эффекты внутримолекулярных взаимодействий: стерические параметры Тафта ( $E_s$  и  $E_s^0$ ), ван-дер-ваальсовы радиусы ( $r_v$ ), молекулярные связьемости ( $\chi$ ) и др. Хотя многие из указанных параметров не отражают в чистом виде только гидрофобные, только электронные или только стерические свойства, приведенное разделение является весьма полезным, например, в связи с проблемой мультиколлинеарности при выборе параметров модели согласно схеме действия биологически активных веществ.

Несмотря на многочисленные факты успешного применения уравнения Хэнча, следует, однако, подчеркнуть, что этот подход и различные его модификации «работают» только в узких рамках рядов родственных соединений. Замена хотя бы одного активного центра в молекуле приводит к изменению характера внутримолекулярных взаимодействий и, как следствие, к нарушению ранее установленных корреляций между физико-химическими параметрами, описывающими соединение, и проявляемым им свойством. В этом отношении, описанные ранее подходы с возможностью отбора общих элементов структуры выглядят предпочтительнее, поскольку допускают изучение широкого спектра соединений. В связи с этим, возникает вопрос: нельзя ли подход Хэнча распространить на анализ соединений из разных рядов и классов путем его объединения с методами выделения общих структурных фрагментов? Впервые эта идея в рамках проблемы ССА была реализована в [43], где принимались во внимание следующие обстоятельства. В традиционном дескрипторном анализе сходства структур используется набор ДЦ с их классификацией по типу атомов и характеру связей. Под влиянием внутримолекулярного окружения, границы, разделяющие электронные свойства ДЦ, могут стать диффузными и это повлечет за собой осложнение в разделении активностей. Отсюда становится очевидным, что в основе классификации ДЦ должен лежать не тип атома, а его способность к межмолекулярному взаимодействию.

Количественной мерой указанного взаимодействия может служить энталпия комплексообразования ( $\Delta H$ ), пропорциональная свободной энергии ( $\Delta G$ ) взаимодействия молекул, содержащих по одному электронодонорному или акцепторному центру [25]. К настоящему времени накоплен обширный экспериментальный материал значений  $\Delta H$ , который позволяет предсказывать термодинамические свойства всего многообразия комбинаций пар донор — акцептор, что дает возможность предложить новую классификацию ДЦ. Для этого необходимо иметь либо экспериментальные данные по  $\Delta H$  для всех возможных донорных молекул с одним и тем же акцептором, либо найти систему эмпирических параметров, позволяющих оценивать  $\Delta H$  взаимодействия любой пары до-

нор — акцептор. У первого варианта (см., например, донорные числа Гутмана [44]) есть существенные ограничения: недостаточная растворимость некоторых комплексов или возможность протекания необратимой реакции. Гораздо перспективнее второй вариант, для реализации которого широко используются два уравнения, предложенные в [45, 46]. В последней модификации формулы [45] энталпия комплексообразования выражается через сумму трех попарных произведений шести типов параметров. Гораздо проще формула, которая в [43] взята за основу классификации донорно-акцепторных взаимодействий.

Согласно [46], энталпия комплексообразования пропорциональна произведению двух безразмерных величин: фактора  $i$ -кислоты (акцептора,  $P_i$ ) и  $j$ -основания (донора,  $E_j$ )

$$\Delta H_{ij} = -\Delta H_{11} P_i E_j, \quad (1)$$

где  $H_{11}$  — коэффициент, определяемый по какому-либо стандартному комплексу. В развитие этого подхода, в [43] предложена единая шкала донорных и акцепторных факторов, которая строится на основе следующей модификации уравнения (1):

$$\Delta H_{ij} = |\Delta H_{11}| E_i E_j,$$

где  $E_i = -P_i$ . В последующих работах [21, 22, 47] этим подходом были охвачены не только нейтральные, но и заряженные доноры и акцепторы. В [48] на основе этой единой шкалы предложена классификация ДЦ для анализа ССА. Список значений  $E_i$  для 70 широко используемых акцепторов и значений  $E_j$  для 230 разнообразных электронодонорных соединений представлен в [22]. Даже этот относительно небольшой массив данных дает возможность рассчитать значения энталпий комплексообразования более чем для 16 тыс. комплексов.

Проведенный анализ больших массивов экспериментальных данных по комплексообразованию позволил установить взаимосвязь между значениями  $E_i(E_j)$  и константами заместителей, характеризующих индукционные влияния и эффект сопряжения [49, 50]. Такой подход потребовал расчета значений этих констант для сложных функциональных групп. В связи с этим, в [52, 53] выведены эмпирические уравнения, описывающие передачу индукционного эффекта через атомы углерода, кислорода и серы. Заключительным этапом внедрения донорно-акцепторных факторов в практику конструирования биологически активных веществ следует считать создание автоматизированной системы оценки значений  $E_i(E_j)$ . Недавно авторами этого обзора предложен и реализован на мини-ЭВМ NORD-10 алгоритм расчета  $E_i(E_j)$  с автоматическим определением констант Тафта сложных функциональных групп по введенной в ЭВМ структурной формуле и описанным типам атомов<sup>1</sup>. Тем самым, создано математическое обеспечение для реализации физико-химического подхода к описанию и классификации соединений по электронным свойствам дескрипторных центров, используемых для отбора элементов структуры, ответственных за конкретный вид биологической активности. Этот подход отличается ясностью и содержательностью легко интерпретируемых принципов классификации, достаточной для широкого использования простотой и уже используется, по крайней мере на качественном уровне, при конструировании веществ с заданными свойствами [54].

Наряду с корректной оценкой электронных свойств ДЦ в проблеме ССА не менее важным является точное описание пространственной структуры молекул с учетом динамики внутримолекулярных движений: таутомерных и конформационных переходов, колебаний большой и малой амплитуд. И здесь можно сразу отметить, что включение в описанную ранее матрицу связности ДЦ в качестве недиагональных элементов значений числа связей между соответствующими ДЦ явно не соответст-

<sup>1</sup> Сапегин А. М., Раевский О. А., Чистяков В. В., Мартынов И. В. //Хим.-фармацевт. журн. 1987. Т. 21, № 9. С. 1098.

звут современным представлениям о пространственной структуре молекул. Включение в матрицу связности равновесных расстояний [17, 33], использование в качестве дескрипторов моментов инерции и молекулярных объемов [21], конечно же, улучшают ситуацию, но не решают проблемы корректного описания пространственной структуры молекул. В качестве простого способа описания молекулярной структуры, довольно полно удовлетворяющего требованиям «идеального», в [55] предложено использовать представление пространственного строения молекулы в виде кривой радиального распределения, как это принято в газовой электронографии. Можно отметить интересную попытку использования аналогичного преобразования молекулярной структуры для разделения транквилизаторов и седативных средств [56], что оказалось намного эффективнее, чем использование на этом же материале фрагментарных дескрипторов [57].

Описанные способы представления структуры молекул в достаточной степени отражают строение соединений в рамках использования ЭВМ для установления связи структура — активность<sup>2</sup>. Причем, выбор формы описания структуры молекул тесно связан с конкретной задачей исследования (классификация соединений по типу проявляемой биоактивности, информационный поиск родственных соединений, выделение общих элементов структуры, количественный прогноз о величине биоактивности и т. д.), а также используемым для этой цели методом. С этой точки зрения, очень важным является вопрос об оптимальном соответствии формы описания структуры и выбранного метода обработки структурной информации.

#### IV. МЕТОДЫ ВЫЯВЛЕНИЯ СВЯЗИ СТРУКТУРА — АКТИВНОСТЬ

Существующие многочисленные методы выявления связи структура — активность по типу решаемых задач можно разделить на две большие группы [21]: методы, направленные на оптимизацию свойств в ряду родственных соединений и методы, ориентированные на поиск общих структурных фрагментов в молекулах соединений, проявляющих определенный вид биоактивности.

##### 1. Оптимизация свойств в ряду родственных соединений

В основе этой группы методов лежат различные многопараметрические регрессионные модели, связывающие структурные параметры соединений с их биоактивностью [58—62]. В табл. 1 представлены основные подходы к обработке структурной информации, использующиеся в ССА для оптимизации свойств в ряду родственных соединений.

Успешное применение описанной в предыдущей главе модели Хэнча способствовало появлению ее многочисленных обобщений и модификаций, представляющих своеобразный «скрининг» разнообразных наборов физико-химических и структурных параметров и типов регрессионных моделей. Уже к 70-ым годам текущего столетия было выведено более 2000 регрессионных зависимостей для более чем 20 000 химических соединений [63], связывающих гидрофобные, электронные, стерические и другие свойства заместителей с проявляемой веществом биологической активностью. Некоторые из этих соотношений были использованы как для прогнозирования эффективности [64—66], так и для выявления механизма действия биологически активных соединений [18, 67]. Однако немало среди них, по выражению Хэнча [68], и «удивительных уродцев». К ним относятся уравнения нереальные по физическому смыслу или неэффективные для прогноза величины биоактивности, но обладающие высокими значениями коэффициентов корреляции.

<sup>2</sup> Raeovsky O. A. // QSAR in Drug Des. Toxicology / Eds Hadzi D. et al. Amsterdam: Elsevier, 1987. P. 31.

Таблица 1

## Основные методы оптимизации свойств в ряду родственных соединений

| Вид анализа                   | Принцип обработки информации  | Область применения   | Ссылки                 |
|-------------------------------|---|--|------------------------|
| Многомерный статистический    | Расчет статистических характеристик распределений случайных величин — средних, дисперсий и т. д.          | Исключение ошибочных данных и анализ «выбросов»  | [72—76]                |
| Многомерное шкалирование      | Сокращение размерности многомерного пространства с целью исключения малоинформативных данных и переменных | Отбор соединений с наибольшей информативностью   | [73—75]                |
| Корреляционный                | Выявление связей между переменными на основе корреляционной матрицы                                       | Проверка гипотезы о независимости параметров модели  | [78, 119]              |
| Факторный                     | «Сжатие» информации путем введения новых переменных, являющихся линейными комбинациями исходных           | Выявление «структурь» данных — сгруппированности объектов и параметров модели                      | [79—82]                |
| Регрессионный и дисперсионный | Построение функциональной зависимости между независимыми переменными и откликом (зависимой переменной)    | Выявление количественной связи между параметрами структуры и проявляемой биоактивностью соединений | [31, 61—68]<br>[85—87] |

Существование таких уравнений показывает, что высокое значение коэффициента корреляции  $R$ , квадрат которого численно равен доле дисперсии функции биоотклика ( $\sigma_y^2 - \sigma^2$ ), «объясненной» регрессионной зависимостью

$$R^2 = (\sigma_y^2 - \sigma^2) / \sigma_y^2,$$

где  $\sigma_y$  — дисперсия функции биоотклика,  $\sigma$  — остаточная дисперсия; является необходимым, но не достаточным условием хорошего качества регрессионной модели. Это объясняется тем, что для регрессионного анализа существует ряд ограничений, нарушение которых приводит к «бедам» регрессионных коэффициентов [69], характеризующих вклад каждого из параметров структуры соединения в величину биоактивности.

Так, в работе [70] обращается внимание на использование без должного обоснования предположения модели Хэнча о нормальном законе распределения для переменных, входящих в эту модель. Анализ 14 возможных законов распределения для переменных в модели Хэнча показал, что независимо от вида характеристики гидрофобности биологически активного соединения наилучшей моделью, описывающей транспорт вещества до биомишени, является показательно-степенной закон распределения величины параметра гидрофобности с явно выраженной асимметричностью. Структура этого распределения свидетельствует о том, что вероятность достижения субстратом места действия возрастает пропорционально его гидрофобности согласно степенному закону, а убывает — по экспоненциальному закону. Наличие асимметрии указывает, что случайные события в данной системе не всегда являются взаимно-независимыми. Наличие корреляций между «независимыми» параметрами является одной из наиболее распространенных причин плохой прогнозической способности выведенных соотношений. Появление неэффективных для прогноза КССА моделей, помимо особенностей «техники» регрессионного анализа, связано и с рядом принципиальных трудностей: а) в установлении однотипности механизма взаимодействия исследуемого набора соединений с одними и теми же рецепторами [1]; б) в выборе регрессионной модели, отвечающей механизму действия вещества; в) в выборе переменных, адекватно описывающих анализируемое свой-

ство; г) в использовании данных по биоактивности, полученных разными исследователями (проблема точности и воспроизводимости биологического эксперимента); д) в выявлении зависимости результирующего биологического действия вещества, влияющего на многие физиологические процессы, происходящие в организме.

Перечисленные трудности, а также то обстоятельство, что в настоящее время не существует альтернативы регрессионному анализу как средству получения количественного прогноза величины биоактивности, приводят к мысли о построении регрессионных моделей после предварительного качественного анализа имеющихся данных. В этом случае последовательность анализа КССА могла бы выглядеть следующим образом. На первом этапе предварительного или, как его еще называют «разведочного»<sup>3</sup>, анализа устанавливается общая структура данных. Для этой цели могут быть использованы как методы наглядного представления многомерного пространства данных, так и автоматические методы классификации и статистической обработки информации. На этом этапе в рамках проблемы структура — активность можно решить две задачи: выделить классы родственных соединений и осуществить переход к набору существенно меньшего числа информативных параметров, описывающих структуру соединений. На следующем этапе может быть проверена гипотеза о независимости параметров модели с помощью методов корреляционного анализа [78]. При необходимости, с использованием методов факторного анализа, в регрессионную модель могут быть введены факторы, представляющие линейную комбинацию исходных параметров модели [79—82]. При формировании окончательной регрессионной модели в [18] рекомендуют придерживаться следующих правил: 1) рассматривать на этапе предварительного анализа как можно больше независимых параметров в качестве переменных моделей; 2) выбирать из различных комбинаций переменных те из них, которые содержат параметры, относящиеся к разным типам взаимодействия (например,  $\lg P$  и  $\sigma$  [83]); 3) отбрасывать переменные с низким уровнем значимости регрессионных коэффициентов; 3) при близких значениях коэффициентов корреляции выбирать модель с наименьшим стандартным отклонением; 5) для исключения случайных корреляций включать по крайней мере 5—6 результатов биологических испытаний на каждый испытываемый в модели параметр [84]; 6) отдавать предпочтение, при прочих равных условиях, более простым моделям; 7) с целью сокращения и упрощения расчетов, а также для избежания принципиальных ошибок учитывать реальность существования модели с точки зрения рассматриваемого явления.

К этому следует добавить и учет возможностей современного регрессионного анализа по оценке качества полученных регрессионных зависимостей [85—87]. Корректное построение регрессионной зависимости позволяет добиться высокой прогностической способности модели, описывающей действие биологически активного вещества, и, тем самым, решить задачу оптимизации свойств в ряду родственных соединений. Однако остается открытый вопрос о прогнозе биоактивности соединений с различающейся структурой. Решение этой проблемы связано с поиском общих элементов структуры.

## 2. Поиск общих элементов структуры биологически активных соединений

В основе поиска общих элементов структуры соединений лежит предположение, что биологическая активность веществ обусловлена наличием в структуре определенных молекулярных подструктур — фармакофоров. Типичным, например, для этого подхода является использование методов распознавания образов (МРО) [87—89] для разделения соединений на активные и неактивные по набору структурных параметров.

<sup>3</sup> Velleman P. F., Hoaglin D. C. Applications, Basics and Computing of Exploratory Data Analysis. Boston: Duxbury Press, 1981. 354 p.

Таблица 2

## Основные методы выявления общих элементов структуры и классификации соединений

| Вид анализа  | Принцип обработки информации   | Область применения   | Ссылки           |
|--|--|--|------------------|
| Дискриминантный  | Построение разделяющих гиперповерхностей по обучающей выборке, состоящей из объектов двух классов  | Отбор информативных параметров модели описания структуры и классификация соединений на активные и неактивные | [14, 77, 97—101] |
| Линейная обучающаяся машина                                | Непараметрический дискриминантный анализ, основанный на исправлении ошибочной классификации через обратную связь в итерационном процессе | То же  | [14, 88]         |
| Логико-структурный   | Вид дискриминантного анализа, основанный на использовании законов индуктивной логики   | »  | [17]             |
| Кластерный   | Классификация объектов по признакам сходства или различия с помощью оптимизации целевой функции  | Классификация соединений по видам активности   | [102—107]        |
| Моделирование пространственной структуры и свойств молекул | Расчет и построение объемных изображений молекул методами квантовой химии и молекулярной механики с помощью молекулярной графики         | Выявление стерической специфичности взаимодействия субстрат — рецептор                                       | [110—116, 119]   |

В этом плане, модель Фри — Вильсона [90], основанная на аддитивном вкладе подструктур заместителей в значение величины биоактивности занимает промежуточное положение между количественными регрессионными моделями и качественными моделями методов распознавания образов. Эта простая модель позволяет существенно сузить область поиска эффективных БАВ. Так, для испытания соединений, где заместители могут находиться в четырех положениях и при изменении их числа в пределах от двух до пяти, применение метода Фри — Вильсона сокращает число необходимых биоиспытаний со 120 до 11 [18]. Однако из-за использования информации на качественном уровне наличия или отсутствия определенных типов фармакофоров в структурной формуле соединения количественный прогноз величины биоактивности часто оказывается ненадежным. Поэтому предпринимались попытки улучшения этой модели, увязывая ее с механизмом действия веществ [91—94] или вводя в рассмотрение взаимное влияние заместителей [95]. Однако в последнем случае число переменных в модели резко возрастает и модель становится труднореализуемой.

Для объектов, описываемых большим числом параметров, для поиска общих элементов структуры более эффективными оказываются методы распознавания образов, основанные на автоматической классификации объектов [96] по сходству или различию. Среди них наиболее широко в литературе по проблеме структура — активность представлены методы, относящиеся к трем видам анализа (табл. 2).

Дискриминантный анализ [97—101] основан на разделении соединений на активные и неактивные путем построения разделяющей гиперплоскости или другой поверхности в пространстве параметров соединений, описывающих их структуру. В рамках этого анализа проводится отбор наиболее информативных, с точки зрения разделения соединений, параметров, строится разделяющая поверхность и проверяется эффективность реализованной дискриминантной модели на обучающей и экзаменационной выборках соединений. Примером успешного применения дискриминантного анализа могут служить работы [99—101].

Логико-структурный анализ [17] использует алгоритмы методов дискриминантного анализа с представлением информации в виде, естественном для исследователя. Многие процедуры обработки информации ориентированы на диалоговый режим работы ЭВМ.

Кластерный анализ [102] состоит в группировке объектов как существий точек в многомерном пространстве, описывающих эти объекты, параметров. Этот вид анализа тесно связан с методами многомерного шкалирования (отображением точек многомерного пространства на поверхность меньшей размерности) и численной таксономии (разбиение объектов на заданные классы). Отличительной чертой методов кластерного анализа является отсутствие пустых кластеров. Методы кластерного анализа в проблеме структура — активность могут использоваться как самостоятельно [104—106], так и в сочетании с другими методами [77, 107, 108].

Во всех рассмотренных методах автоматической классификации топологические или топографические особенности строения молекул отображались в виде набора индексированных параметров, — геометрических дескрипторов, являющихся координатами соединений, как точек в многомерном пространстве параметров. С развитием компьютерной графики в последние годы появилась возможность наглядного трехмерного изображения самих структур молекул на экране графического дисплея ЭВМ [108—110]. Это положило начало еще одному направлению в исследованиях по ССА — молекулярному моделированию пространственной структуры соединений.

Наиболее плодотворным оказалось использование методов моделирования трехмерной структуры молекул для изучения пространственной специфичности взаимодействия субстрат — реагент по принципу «ключ — замок» [110], особенно, при представлении электронной «формы» молекулы как поверхности вокруг ядерного скелета конформации молекулы [27], полученной с помощью квантовохимического расчета [111] или метода молекулярной механики [36]. Этот подход позволяет обнаружить сходство в электронно-топографических свойствах молекул, не очевидное при других методах анализа. Количественная сторона учета стерической специфичности взаимодействия субстрат — реагента разрабатывалась в регрессионном анализе [112—115], однако, в последнее время наметилась тенденция к совместному использованию методов молекулярного моделирования и регрессионного анализа [115, 116], которая, в перспективе, должна привести к системному подходу к проблеме связи структура — активность.

### 3. Системный подход к анализу связи структура — активность

Идея системного подхода к анализу ССА состоит в создании единой модели поиска биологически активных соединений, охватывающей как качественную и количественную, так и априорную (экспертную) информацию об изучаемом классе соединений в рамках экспертной системы, включающей соответствующий набор методов выявления ССА. Целесообразность этого продиктована как необходимостью учета всей имеющейся у исследователя информации, так и выбором оптимального метода для ее обработки. Действительно, если, например, сформулировать проблему связывания субстратом рецептора как задачу нахождения функции плотности распределения вероятности комплексообразования по набору параметров, характеризующих структуру субстрата, то станет очевидной связь этой проблемы с задачей восстановления зависимостей в условиях обратной некорректной задачи [87]. Из теории решения подобных задач [117] можно заключить, что качество полученного результата в значительной степени зависит от полноты учета специфики конкретной задачи, выбора соответствующего метода ее решения. Поскольку на разных уровнях и стадиях анализа ССА исследователь располагает информацией, отличающейся как по количеству и качеству, так и по форме ее представления и информативности, то становится по-

нятной необходимость использования не какого-то одного метода выявления ССА, а всего комплекса этих методов, причем в такой их последовательности, которая, в конечном счете, привела бы к количественному прогнозу величины биологической активности неизученного соединения по параметрам, характеризующим его структуру.

Первым шагом в направлении объединения качественных и количественных методов анализа ССА в рамках единой модели стала разработка физико-химического подхода, подробно описанного в гл. III.

Возможность системного анализа связи структура—активность в этом подходе связана с унификацией описания электронно-топографических свойств молекул на базе единой донорно-акцепторной модели взаимодействия субстрат—рецептор, где параметры модели могут выступать как индикаторные переменные на этапе автоматической классификации соединений и как количественные переменные при восстановлении регрессионных зависимостей структура—активность. Это позволяет ставить вопрос об оптимизации свойств в ряду соединений с отличающимися структурными формулами, в том числе принадлежащих разным классам, но имеющих близкие донорно-акцепторные и топографические свойства дескрипторных центров.

Современная тенденция к интеграции методов и подходов в рамках проблемы ССА нашла свое отражение и в разрабатываемых в настоящее время комплексах программ.

## **V. КОМПЛЕКСЫ ПРОГРАММ ПО ВЫЯВЛЕНИЮ СВЯЗИ СТРУКТУРА — АКТИВНОСТЬ**

Рассмотренные методы и подходы к анализу ССА для своей реализации требуют создания соответствующих программных комплексов. Существующие комплексы (пакеты) программ различаются как по назначению и степени охвата решаемой проблемы, так и по ориентации на определенный класс ЭВМ. Наиболее удобными являются пакеты программ, организованные по модульному принципу с диалоговым режимом работы. Как правило, такие комплексы ориентированы на мини-ЭВМ и, благодаря легкости включения новых программ в комплекс, обладают несомненным преимуществом перед другими пакетами программ при создании экспертной системы по проблеме ССА на базе лабораторной мини-ЭВМ.

В качестве примера, рассмотрим четыре таких комплекса, в которых реализованы описанные ранее подходы к анализу ССА: автоматической классификации на основе методов распознавания образов (ADAPT) логико-структурный (ТОПЛОГ), молекулярного моделирования (MOLY) и физико-химический (РАСТР) подходы.

В пакете программ ADAPT [14] основной обрабатываемой информацией являются структурные формулы соединений. Любая из них может быть введена в ЭВМ путем построения ее изображения на экране графического дисплея с помощью специального набора команд в интерактивном режиме. Введенные структурные формулы хранятся в виде матриц связности на дисковых накопителях. Предусмотрены возможности для добавления, исключения, изменения и вызова хранимых структур соединений. Анализ ССА проводится на основе методов параметрической и непараметрической статистики и распознавания образов: линейной регрессии, дискриминантного анализа, линейной обучающейся машины и классификации по количеству ближайших соседей.

Система ADAPT реализована на базе 16-ти битовой мини-ЭВМ MODCOMP II/25 и вместе с банком данных занимает около 2 Мбайт памяти на магнитных дисках.

Программный комплекс ТОПЛОГ [17] представляет пользователю более развитую систему ввода, редактирования и преобразования информации к виду, естественному для восприятия исследователем. Это нашло свое отражение в файловой системе комплекса, насчитывающего 14 типов файлов различного назначения. В комплексе реализованы про-

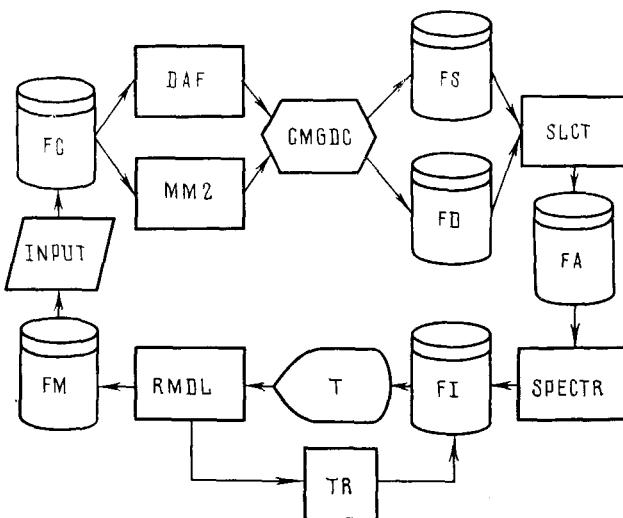


Рис. Блок-схема основного ядра комплекса РАСТР.

Программы: INPUT — ввода, DAF — расчета дононно-акцепторных факторов, MM2 — расчета равновесных конформаций молекул методом молекулярной механики, CMGDC — формирования МГДЦ, SLCT — отбора общих структурных фрагментов, SPECTR — построения спектра межатомных расстояний, RMDL — построения регрессионных моделей, TR — преобразования регрессионных моделей фармакофоров в спектр межатомных расстояний; файлы: FC — соединений, FS — структур, FD — структур банка данных, FA — активностей (фармакофоров), FI — изображений (спектров межатомных расстояний), FM — регрессионных моделей

цедуры логико-структурного подхода, подробно описанного в [17]. Отличительной особенностью пакета ТОПЛОГ, является наличие алгоритма отбора общих структурных фрагментов молекул на основе анализа граф-представлений структурных формул исследуемых соединений. Пакет состоит из 14 программ общим объемом около 8 тыс. фортрановских операторов и используется на мини-ЭВМ НР-1000.

В отличие от описанных комплексов, где основным видом информации является структурная формула соединений, в комплексе MOLY [10, 118] для конструирования БАС используется объемное трехмерное изображение молекулы с отображением ее свойств (например, поверхности электростатического потенциала). Это позволяет, с помощью методов молекулярного моделирования, получать параметры структуры молекулы, отражающие пространственную специфичность взаимодействия субстратов с рецепторами, которые могут быть использованы для построения регрессионных моделей КССА. Комплекс MOLY содержит около 15 тыс. фортрановских операторов и занимает объем памяти, равный 1 Мбайту.

В комплекс программ РАСТР (распознавание активных структур)<sup>4</sup> включены процедуры, реализующие физико-химический подход к выявлению общих по электронно-топографическим свойствам фрагментов БАВ в рамках проблемы структура — активность. Отличительной чертой комплекса является возможность системного подхода к анализу ССА с использованием на разных этапах исследования всего набора рассмотренных ранее методов (табл. 1 и 2). Это достигается унифицированным описанием электронно-топографических свойств молекул с позиции единой дононно-акцепторной модели взаимодействия субстрат — рецептор. Так, например, одним из вариантов работы с пакетом РАСТР может быть следующий. С терминала ЭВМ вводится структурная формула соединения и информация о типе входящих в нее атомов. С помощью программы INPUT (рисунок) эти данные преобразуются в матрицу связности и записываются в файл соединений (FC) на магнитном диске. Информация из этого файла используется для расчета матрицы геометрии дескрипторных центров (МГДЦ) — физико-химического «портрета»

<sup>4</sup> Раевский О. А., Сапегин А. М./Хим.-фармацевт. журн. 1987. Т. 21. С. 1338; Сапегин А. М., Раздольский А. Н., Чистяков В. В., Раевский О. А./Там же. С. 1341.

молекулы, описывающего донорно-акцепторные и топографические свойства активных центров субстратов. Диагональные элементы МГДЦ кодируются на основе рассчитанных значений донорно-акцепторных факторов (ДАФ) ДЦ в соответствии с унифицированной шкалой донорно-акцепторных взаимодействий. Алгоритм расчета ДАФ ДЦ (программа DAF) построен на анализе дерева граф-представления структурной формулы соединения с использованием банка данных по константам Тафта для различных типов атомов и корреляционных уравнений для расчета констант Тафта сложных заместителей и ДАФ ДЦ [50, 52–53]. Внедиагональные элементы МГДЦ содержат информацию о топографии ДЦ в виде расстояний между соответствующими ДЦ для равновесных конформаций, рассчитанных методом молекулярной механики [36] (программа MM2). Матрица геометрии ДЦ формируется программой CMGDC, которая работает в двух режимах: анализа ССА и режиме записи МГДЦ в банк данных.

С помощью программы SLCT отбираются общие структурные фрагменты для исследуемого (файл FS) и хранимых в банке данных (файл FD) соединений. Процедура поиска общих структурных фрагментов аналогична используемой в комплексе ТОПЛОГ [24]. Отличие состоит лишь в том, что в данном случае под общим структурным фрагментом понимается не строгая тождественность элементов структурной формулы соединений, а их регулируемая исследователем близость по электронно-топографическим свойствам. Выделенные фрагменты являются потенциальными фармакофорами и записываются в файл биологически активных структур (файл FA).

Содержащиеся в файле FA МГДЦ общих структурных фрагментов БАС могут быть представлены (с помощью программы SPECTR) в виде спектра межатомных расстояний, удобном для визуального сравнения структур соединений, или набора структурных параметров для построения регрессионных моделей ССА. Спектры межатомных расстояний хранятся в файле «изображений» (FI).

Таким образом, на описанном выше этапе, решается качественная задача по выявлению потенциальных фармакофоров, т. е. на основе индикаторных параметров донорно-акцепторной модели формируется представление об электронных и пространственных свойствах активных центров субстратов, ответственных за конкретный вид биологической активности.

Следующим важным шагом является установление количественной зависимости между параметрами структуры соединения и величиной биоотклика. В рассматриваемом комплексе РАСТР исходной информацией для построения регрессионных моделей (программа RMDL) является информация из файла FI. Поиск оптимальной регрессионной зависимости осуществляется на основе принципа структурной минимизации среднего эмпирического риска [87]. Оценка прогностической способности модели проводится по экзаменационной выборке данных.

Полученные на этапе регрессионного анализа результаты могут быть повторно использованы в комплексе РАСТР для уточнения модели или построения новой. Эти данные хранятся в файле FM. Благодаря файловой организации передачи информации между программами комплекса, он обладает большой гибкостью по отношению к включению стандартных или оригинальных методов обработки информации на любом этапе исследования и может аккумулировать в себе все виды анализа КССА, став основой для построения развитой экспертной системы.

Все программы комплекса написаны на языке ФОРТРАН-77 и используются на мини-ЭВМ NORD-10 (емкость ОЗУ 128 Кбайт).

Рассмотренные комплексы программ показывают, что к настоящему времени программное обеспечение охватывает широкий круг алгоритмов обработки структурной информации, создавая, тем самым, базу для системного подхода к анализу ССА. Практической реализацией этой идеи могла бы стать разработка проблемно-ориентированной экспертной системы по конструированию веществ с заданными свойствами.

## VI. ЗАКЛЮЧЕНИЕ

Для поиска эффективных биологически активных синтетических соединений на базе основополагающих принципов конструирования веществ с заданными свойствами необходимо решение следующих задач: 1) идентификация соединения, взаимодействующего с рецептором; 2) установление электронной и пространственной структуры этого соединения; 3) корректное описание его структуры в рамках использования ЭВМ; 4) получение статистически достоверных данных по биологической активности соединений; 5) выбор оптимального метода или группы методов для решения конкретной задачи по выявлению связи структура—активность.

Некорректный подход к решению любой из указанных задач может помешать установлению «истинной» взаимосвязи структура—активность или даже привести к установлению ложных закономерностей. Так, например, отождествление субстрата с введенным в организм веществом (т. е., пренебрежение возможностью химических и биохимических трансформаций вещества в процессе транспорта от места введения до биомишени) подчас затрудняет выявление общих фрагментов структуры среди соединений, проявляющих один и тот же вид биоактивности. Загрубление модели описания электронной и пространственной структуры соединений автоматически скажется на прогностической способности такой модели. Эффективность же последней будет определяться также выбором метода обработки информации, т. е. его соответствия форме представления и уровню информации о структуре и свойствах молекул.

В литературе уже имеются многочисленные примеры правильного предсказания биологической активности соединений (табл. 3). [119]. Однако подавляющее большинство этих работ связано с оптимизацией свойств в ряду родственных соединений. Качественно же новый уровень в проблеме КССА, способный привести к выявлению принципиально новых классов эффективных препаратов, возможен лишь на основе использования методов отбора общих структурных фрагментов среди разнообразных соединений с последующей оптимизацией проявляемых ими полезных свойств.

Это приводит к необходимости развития системного подхода к проблеме ССА, который должен включать в себя, помимо рассмотренных ранее вопросов, обширный раздел химической информатики [120] и искусственного интеллекта [121]. Завершающим этапом системного исследования ССА могло бы стать планирование синтеза и технологических схем получения новых препаратов [122, 123]. Реализация такого подхода, в конечном счете, должна превратить эмпирический путь установления связи структура—активность в фундаментальное научное направление, аккумулирующее в себе достижения химии, физики, биологии и информатики. Тем самым, будет создана мощная платформа экономически эффективного конструирования веществ с заданными свойствами для нужд промышленности, сельского хозяйства и медицины.

\* \* \*

За время подготовки обзора к печати, появилось значительное число работ, затрагивающих практически все рассмотренные аспекты проблемы целенаправленного поиска и конструирования эффективных БАВ. В первую очередь следует отметить публикацию трудов шестого европейского симпозиума по КССА [124], а также монографию [125], в которой изложены основы построения систем определения основных видов биоактивности для больших массивов химических соединений.

Особенно интенсивно развивается направление, связанное с трехмерным представлением молекулярной структуры методами молекулярной графики [126—131], основанным на эмпирических [132, 133] и квантово-химических [134, 135] расчетах конформаций молекул или соответствующих рентгеноструктурных данных [136, 137]. Много внимания уделяется совершенствованию моделей описания структуры с помощью

Таблица 3

## Примеры правильных предсказаний связи структура—активность [119]

| Тип соединения                             | Вид биоактивности                         |
|--|---|
| Бензотиадиазины                            | Гипотензивное действие                    |
| Пропениламины                              | Ингибированиеmonoаминоксидазы             |
| N-(феноксиэтил)-циклогексил-амины          | То же                                     |
| β-Карболины                                | »   |
| Сульфонамиды                               | Ингибиование карбоангидразы               |
| Карбеноилпептидины                         | Ингибиование холинэстеразы                |
| Кетоны                                     | Ингибиование хинотрипсина                 |
| Бензилпиридиневые ионы                     | Ингибирование комплемента                 |
| Стероиды                                   | Связывание с прогестероновым рецептором   |
| Эфиры                                      | Гидролиз папаина                          |
| Допаминовые эфиры                          | Гидролиз ариламида                        |
| Пептидил- <i>p</i> -нитроанилиды           | Субтилизиновый гидролиз                   |
| Полициклические ароматические углеводороды | Гидроксилирование микросомальной оксидазы |
| Изатип-β-изотиосеникарбазоны               | Антивирусная канцерогенность              |
| Пиразолы                                   | То же                                     |
| Производные цитидина                       | Цитостатическая, иммунодепрессивная       |
| Нитрозоуретаны                             | Цитостатическая                           |
| Митониины                                  | »   |
| Производные гиданттоина                    | »   |
| Рафинициновые амины                        | Антибактериальная                         |
| Сульфонамиды                               | »   |
| Эритромицины                               | »   |
| Хинохалин-1,4-диоксиды                     | »   |
| Фенантрены                                 | Антимикоплазматическая                    |
| Фениламиноуретаны                          | »   |
| Нафтохионы                                 | »   |
| Туберины                                   | Антимикоплазматическая                    |
| Метоксихлорины                             | Инсектицидная                             |
| N-гидроксисиридоны                         | Фунгицидная                               |
| Производные пиридина                       | »   |
| Трифторметансульфонамиды                   | Гербицидная                               |
| Алкиларилкарбонильные соединения           | »   |
| Ароматические эфиры                        | »   |
| Пиперидиноацетанилиды                      | Фитоактивность                            |
| Пиримидоны                                 | Ингибирование реакции Хилла               |
| Уретаны                                    | Гербицидная                               |
| Фенилацетамиды                             | Гипнотическая                             |
| Пиридинметаны                              | Спазмолитическая                          |
| Проназины                                  | Нейролептическая                          |
| Ароматические и алифатические кислоты      | Противовоспалительная                     |
| Производные бензотиепина                   | Нейролептическая                          |
| Карбоксилатные эфиры                       | Противотромбическая                       |
| Акрилаты                                   | Токсичность                               |
| Медные хелаты                              | Цитотоксичность                           |
| Пираненамины                               | Иммунодепрессивная                        |
| Азапурин-6-оны                             | »   |
| Фенилоксазолидины                          | Радиозащитная                             |
| Карбоноилэтиламины                         | Холинолитическая                          |
| Триазины                                   | Ингибирование дегидрофолатредуктазы       |
| Бенздиазепины                              | Антителентетразольная активность          |
| Цианоциклогексиламины                      | Фунгицидная                               |
| Пиперидиноацетанилиды                      | Антисеротониновая                         |
| Аналоги соматостатина                      | Антидиабетическая                         |
| Стероиды                                   | Прогестероновая                           |
| Азациклоалканы                             | Анальгетическая                           |
| Органокремниевые амины                     | Фунгицидная                               |
| Индадионы-1,3                              | Токсичность                               |
| Сульfonyны                                 | Антибактериальная                         |
| Алкобензилалкиламины                       | Токсичность                               |
| Спиробарбитуровые кислоты                  | Седативная                                |
| Аналоги аденоцина                          | Гипотензивная                             |
| Аналоги триметоприма                       | Антибактериальная                         |

наборов дескрипторов [138—149] и наглядному отображению результатов статистического анализа структурной информации [150—153].

К числу принципиально новых тенденций в проблеме структура — активность, обозначившимся в последнее время, следует отнести углубление физико-химической содержательности моделей КССА [154—157], включение в них, наряду с фармакодинамической, фармакокинетической стадии [158—164] действия БАВ и революционизирующее влияние технологии программирования в экспертных системах [165—171].

Создание экспертной системы КССА могло бы стать основой для формирования своеобразной «программной» теории целенаправленного поиска и конструирования биологически активных веществ.

## ЛИТЕРАТУРА

1. Хэнч К.//Хим.-фармацевт. журн. 1980. Т. 14. № 10. С. 15.
2. Glasser A. S.//Meth. Find.: Exptl. Clin. Pharmacol. 1984. V. 6. N 10. P. 563.
3. Балобанов А. Е. Развитие количественных представлений о связи между структурой и биологической активностью химических соединений. М.: Ин-т истории естествознания и техники, 1984. Рук. деп. ВИНИТИ № 1639-84.
4. Blake J.//Arch. Gen. Med. III Ser. 1839. V. 6. P. 289.
5. Hansch C.//Drug Metabolism Rev. 1984—85. V. 15. N 7. P. 1270.
6. Richardson B. W.//Brit. Assoc. Rep. 1864. V. 34. P. 120; 1867. V. 37. P. 47; 1868. V. 38. P. 179.
7. Crum-Brown R., Fraser T.//Trans. Roy. Soc. 1869. V. 25. P. 151; 1869. V. 25. P. 693.
8. Ariens E. J., Simonis A. M.//Top. Curr. Chem. 1974. V. 52. P. 1.
9. Ландай М. А. Молекулярные механизмы действия физиологически активных соединений. М.: Наука, 1981.
10. Computer-Assisted Drug Design./Eds Olson E. S. et al. Washington: Amer. Chem. Soc, 1979.
11. Strategy in Drug Design./Ed. Buisman I. A. K. Amsterdam: Elsevier, 1982.
12. Quantitative Structure-Activity Relationships of Drugs./Ed. Topliss J. G. N. Y.: Acad. Press, 1983.
13. QSAR and Strategies in Design of Bioactive Compounds./Ed. Seydel J. K., N. Y.: UCH Publ., 1985. 442 p.
14. Стьюпер Э., Брюггер Й., Джурс П. Машинный анализ связи химической структуры биологической активности/Пер. с англ. М.: Мир, 1982. 235 с.
15. Drug Design./Ed. Ariens E. J. N. Y.: Acad. Press, 1979. 420 p.
16. Голендер В. Е., Розенблит А. Е. Вычислительные методы конструирования лекарств. Рига: Зинатне, 1978. 232 с.
17. Розенблит А. Б., Голендер В. Е. Логико-комбинаторные методы в конструировании лекарств. Рига: Зинатне, 1983. 352 с.
18. Нижний С. В., Эпштейн Н. А.//Успехи химии. 1978. Т. 47. № 4. С. 739.
19. Овчинников А. А., Болдырев А. И.//Там же. 1986. Т. 55. С. 539.
20. Валуева Л. Н., Засецин В. М., Промоненков В. К.//Применение математических методов для анализа связи молекулярная структура — пестицидная активность. Сер. Химические средства защиты растений. М.: НИИТЭХИМ, 1985. 56 с.
21. Раевский О. А. Введение в конструирование биологически активных веществ. М.: МХТИ, 1984. 80 с.
22. Раевский О. А., Григорьев В. Ю. Задачи по курсу «Конструирование биологически активных веществ». М.: МХТИ, 1986. 80 с.
23. Simon Z.//Angew. Chem. Int. Ed. 1974. V. 13. № 11. P. 719.
24. Hayward R. C.//Chem. Soc. Rev. 1983. V. 12. № 3. P. 285.
25. Раевский О. А.//Теорет. эксперим. химия. 1986. № 4. С. 450.
26. Крицман В. А.//Журн. Всесоюз. хим. о-ва им. Д. И. Менделеева. 1975. Т. 20. № 6. С. 612.
27. Lehn J. M.//Recherche. 1981. N 127. P. 1213.
28. Kellogg R. M.//Ibid. 1984. N 156. P. 819.
29. Fischer E.//Chem. Ber. 1894. B. 27. N 3. S. 2985.
30. Берцов Т. Т., Коровкин Б. Ф./Биологическая химия. М.: Медицина, 1983. С. 137.
31. Hansch C., Fujita T.//J. Amer. Chem. Soc. 1964. V. 86. P. 1616.
32. Авидон В. В., Лексина Л. А.//НТИ. Сер. 2. 1974. № 3. С. 22.
33. Гитлина Л. С., Голендер В. Е., Дробоглав В. В. и др. Методы представления и обработки структурной информации для анализа связи структура — активность. Препринт. Рига: Ин-т орг. синтеза АН ЛатвССР, 1981. 74 с.
34. Авидон В. В., Козлова С. П., Ароолович В. С.//НТИ. Сер. 2. 1974. № 12. С. 21.
35. Wiswesser W. J. A Line-Formula Notation. N. Y.: T. J. Cro. Co, 1954.
36. Буркет У., Эллинджер Н. Молекулярная механика/Пер. с англ. М.: Мир, 1986. 364 с.
37. Kier L. B., Hall L. H. Molecular Connectivity in Chemistry and Design. N. Y.: Acad. Press, 1976. 257 p.
38. Kaufman J. J.//Int. Quant. Chem. 1979. V. 26. P. 221.
39. Лужков В. Б., Богданов Г. Н.//Успехи химии. 1986. Т. 55. № 1. С. 3.

40. Rollett J. S. //J. Mol. Struct. 1985. V. 130. P. 1.  
 41. Димоголо А. С. //Хим.-фармацевт. журн. 1985. № 3. С. 438.  
 42. Motoc I. //Lecture Notes in Chemistry. 1980. V. 15. P. 2.  
 43. Раевский О. А., Новиков В. П. //Хим.-фармацевт. журн. 1982. № 5. С. 583.  
 44. Бургер К. Сольватация, ионные реакции и комплексообразование в неводных растворах/Пер. с англ. М.: Мир, 1984. 256 с.  
 45. Kroeger M. K., Drago R. S. //J. Amer. Chem. Soc. 1983. V. 103. P. 3250.  
 46. Иогансен А. В. //Теорет. и эксперим. химия. 1971. № 7. С. 302, 312.  
 47. Раевский О. А., Григорьев В. Ю., Соловьев В. П. //Хим.-фармацевт. журн. 1984. № 5. С. 578.  
 48. Раевский О. А., Авидон В. В., Новиков В. П. //Там же. 1982. № 8. С. 968.  
 49. Мартынов И. В., Раевский О. А. //Вестн. АН СССР. 1983. № 7. С. 99.  
 50. Мартынов И. В., Раевский О. А. //Докл. АН СССР. 1982. Т. 265. № 3. С. 664.  
 51. Мартынов И. В., Раевский О. А. //Журн. Всесоюз. хим. о-ва им. Д. И. Менделеева. 1983. Т. 28. № 6. С. 716.  
 52. Мартынов И. В., Раевский О. А. //Докл. АН СССР. 1980. Т. 254. № 6. С. 1404.  
 53. Мартынов И. В., Раевский О. А. //Там же. 1985. Т. 283. № 2. С. 417.  
 54. Тюрина Л. А., Семенов В. А., Кадыров Ч. Ш. и др. //Агрохимия. 1986. № 1. С. 99.  
 55. Новиков В. П., Раевский О. А. //Хим.-фармацевт. журн. 1982. № 5. С. 574.  
 56. Soltzberg L. J., Wilkins C. L. //J. Amer. Chem. Soc. 1977. V. 99. P. 439.  
 57. Stuper A. J., Jurs P. C. //Ibid. 1975. V. 97. P. 182.  
 58. Kubinyi H. //Arzneim-Forsch. 1979. V. 29. N 8. P. 1067.  
 59. Hansch C., Clayton J. M. //J. Pharm. Sci. 1973. V. 62. P. 1.  
 60. Szentpaly L. //J. Amer. Chem. Soc. 1984. V. 106. P. 6021.  
 61. Hansch C. //Drug Intell. Clin. Pharm. 1982. V. 16. N 5. P. 391.  
 62. Mager P. P. //Med. Res. Rev. 1983. V. 3. N 4. P. 435.  
 63. Hansch C. //J. Med. Chem. 1976. V. 19. P. 1.  
 64. Dekker W. H., Selling H. A., Overeem J. C. //J. Agr. Food Chem. 1975. V. 23. P. 785.  
 65. Pursell W. P. //Eur. J. Med. Chem. 1975. V. 10. P. 335.  
 66. Goodford P. G., Norrington F. E. //Quant. Struct.-Act. Relat./Ed. Tichy M., Budapest: Akad. Kiado, 1979. P. 79.  
 67. Marchewka V. C., Ebner N. A., Sehon K. D., Hanna P. E. //J. Med. Chem. 1985. V. 28. N 1. P. 18.  
 68. Montgomery J. A., Mayo J. G., Hansch C. //J. Med. Chem. 1974. V. 17. P. 477.  
 69. Мостеллер Ф., Тьюки Дж. Анализ данных и регрессия/Пер. с англ. М.: Финансы и статистика, 1982. Ч. 1, 2.  
 70. Фрумин Г. Т., Мухоморов В. К. //Хим.-фармацевт. журн. 1985. № 3. С. 180.  
 71. Barth A., Franke R., Orlick M. //Quant. Struct.-Activity Analysis. //Eds Franke R. et al. Berlin: Akad.-Verlag, 1978. P. 395.  
 72. Dillon W. P., Goldstein M. //Multivariate Analysis. Methods and Applications. N. Y.: Wiley, 1984.  
 73. Dove S., Streich W. J., Franke R. //J. Med. Chem. 1980. V. 23. P. 1456.  
 74. Dove S., Coats E., Scharfenberg P., Franke R. //Ibid. 1985. V. 28. N 4. P. 447.  
 75. Streich W. J., Dove S., Franke R. //Ibid. 1980. V. 23. P. 1452.  
 76. Wootton R. //Ibid. 1983. V. 26. P. 275.  
 77. Wold S., Dunn W. J. //J. Chem. Inf. Comput. Sci. 1983. V. 23. N 1. P. 6.  
 78. Жданов Ю. А., Минкин В. И. Корреляционный анализ в органической химии. Ростов-н/Д: РГУ, 1966.  
 79. Иберла К. Факторный анализ/Пер. с англ. М.: Статистика, 1980.  
 80. Weiner M. L., Weiner P. H. //J. Med. Chem. 1973. V. 16. P. 655.  
 81. Franke R. //Farmaco. 1978. V. 34. P. 545.  
 82. Franke R., Dove S., Kühne R. //Eur. Med. Chem. 1979. V. 14. P. 363.  
 83. Craig P. N. //J. Med. Chem. 1971. V. 14. P. 680.  
 84. Topliss J. G., Costello R. J. //Ibid. 1972. V. 15. P. 1066.  
 85. Себер Дж. Линейный регрессионный анализ/Пер. с англ. М.: Мир, 1980. 456 с.  
 86. Moriguchi I., Komatsu K., Matsushita Y. //Quant. Struct.-Act. Relat. 1984. V. 3. N 3. P. 106.  
 87. Алгоритмы и программы восстановления зависимостей/Под ред. Вапника В. Н. М.: Наука, 1984. 816 с.  
 88. Джурс П., Айзенауэр Т. Распознавание образов в химии/Пер. с англ. М.: Мир, 1977. 230 с.  
 89. Henry D. R., Jurs P. C., Denny W. A. //J. Med. Chem. 1982. V. 25. P. 899.  
 90. Free S. M., Wilson J. W. //J. Med. Chem. 1964. V. 7. P. 395.  
 91. Fujita T., Ban T. //Ibid. 1971. V. 14. P. 148.  
 92. Cammarata A., Yau S. J. //Ibid. 1970. V. 13. P. 93.  
 93. Clayton J. M., Pursell W. P. //Ibid. 1969. V. 12. P. 1087.  
 94. Craig P. N. //Ibid. 1972. V. 15. P. 144.  
 95. Kopecky J., Bocek K. //Experientia. 1967. V. 23. P. 125.  
 96. Айвазян С. А., Бежаева З. И., Староверов О. В. Классификация многомерных наблюдений. М.: Статистика, 1974. 238 с.  
 97. Martin Y. C., Holland J. B., Jarboe C. H., Plotnikoff N. //J. Med. Chem. 1974. V. 17. P. 409.  
 98. Henry D. R., Block J. H. //Eur. Med. Chem. 1980. V. 15. P. 133.  
 99. Нигматуллин Р. С., Осипов А. Л., Пузаткин А. П., Контюг В. А. Хим.-фармацевт. журн. 1985. № 2. С. 179.

100. Jurs P. C., Stouch T. R., Czerwinski M., Narvaez J. N.//J. Chem. Inf. Comput. Sci. 1985. V. 25. P. 296.
101. Шульман В. С.//Хим.-фармацевт. журн. 1984. № 9. С. 1087.
102. Дюран Б., Оддел П. Кластерный анализ/Пер. с англ. М.: Статистика, 1977. 128 с.
103. Hansch C., Unger S. H., Forsythe A. B.//J. Med. Chem. 1973. V. 16. P. 1217.
104. Adamson G. N., Bawden D.//J. Chem. Inf. Comput. Sci. 1981. V. 21. P. 204.
105. Wold S.//Pattern Recognition. 1976. V. 8. P. 127.
106. Willett P.//J. Chem. Inf. Comput. Sci. 1984. V. 24. № 1. P. 29.
107. Randic M.//Int. Quan. Chem. Quan. Biol. 1984. N 11. P. 137.
108. Advances in Drug Research. V. 14/Ed. Testa B. N. Y.: Acad. Press, 1985. 339 p.
109. Andy M.//Comput. Graph. Int. Conf. Proc. L.: Pinner, 1983. P. 435.
110. Macromolecular Structure and Specificity: Computer Assisted Modelling and Applications/Eds Kataraghaven B. et al., N. Y.: Acad. Sci., 1985. 209 p.
111. Hopfinger A. J./Pharm. Int. 1984. V. 5. N 9. P. 224.
112. Chose A. K., Crippen G. M.//J. Med. Chem. 1985. V. 28. P. 333.
113. Beddel C. R.//Chem. Soc. Rev. 1984. V. 13. N 3. P. 279.
114. Hopfinger A. J.//J. Med. Chem. 1983. V. 26. P. 990.
115. Selassie C. D., Fang Z.-X., Li R.-L., Hansch C.//J. Med. Chem. 1986. V. 29. P. 621.
116. Кузьмин В. Е., Крутцус С. В./Хим.-фармацевт. журн. 1986. № 7. С. 791.
117. Тихонов А. Н., Гончарский А. В., Степанов В. В., Ягода А. Г. Регуляризирующие алгоритмы и априорная информация. М.: Наука, 1983. 200 с.
118. Dyoit T. M., Stuper A. J., Zander G. S.//J. Chem. Inf. and Comput. Sci. 1980. V. 20. N 1. P. 28.
119. Franke R.//Theoretical Drug Design Methods. Berlin: Akad.-Verlag. 1984. 415 p.
120. Rush J. E.//J. Chem. Inf. Comput. Sci. 1985. V. 89. P. 4426.
121. Clementi E. J.//Phys. Chem. 1985. V. 89. P. 4426.
122. Johnson A. P.//Chem. in Brit. 1985. V. 21. N 1. P. 59.
123. Зефиров М. С., Гордеева Е. В./Успехи химии. 1987. Т. 56. С. 1753.
124. QSAR in Drug Design and Toxicology./Eds Hadzi D. et al. Amsterdam: Elsevier, 1987. 375 р.
125. Баренбойм Г. М., Маленков А. Г. Биологически активные вещества. Новые принципы поиска. М.: Наука, 1986. 363 с.
126. Liebman M. N./J. Mol. Graphics. 1986. V. 4. N 1. P. 61.
127. Ripka W. C.//Nature. 1986. V. 321. P. 93.
128. Rose G. D., Lee R. H.//Biophys. J. 1986. V. 49. N 1. P. 83.
129. White D. N. J., Pearson J. E.//J. Mol. Graphics. 1986. V. 4. N 3. P. 134.
130. Hansch C., Klein T. E.//Acc. Chem. Res. 1986. V. 19. N 12. P. 392.
131. Meyer E. F., Bode W.//QSAR in Drug Des. Toxicol./Eds Hadzi D. et al. Amsterdam: Elsevier, 1987. P. 247.
132. Tomasi J.//Ibid. 1987. P. 269.
133. Labanowski J., Motoc I., Naylor C. B., et al.//QSAR. 1986. V. 5. P. 138.
134. Cummins P. L., Rendell A. P. L., Swanton D. J. et al.//Int. Rev. Phys. Chem. 1986. V. 5. N 2/3. P. 139.
135. Kikuchi Q.//QSAR. 1987. V. 6. P. 179.
136. Fujita T.//Ibid. 1987. V. 6. N 2. P. 54.
137. Hol W. G. J.//Angew. Chem. Int. Ed. Engl. 1986. V. 25. N 9. P. 767.
138. Kier L. B., Hall L. H. Molecular Connectivity in Structure-Activity Analysis. N. Y.: Wiley, 1986. 262 р.
139. Kier L. B.//QSAR. 1986. V. 5. N 1. P. 7.
140. Kier L. B.//Ibid. 1987. V. 6. N 1. P. 8.
141. Kier L. B.//Ibid. N 3. P. 117.
142. Verloop A. The STREIMOL Approach to Drug Design. N. Y.: Marcel Dekker, 1987.
143. Kramer C. R.//Z. Phys. Chem. 1986. V. 267. N 2. P. 277.
144. Trinajstic N., Klein D. J., Randic M.//Int. J. Quant. Chem. Quant. Symp. 1986. V. 20. P. 699.
145. Dunn W. J. III., Wold S.//Trends Anal. Chem. 1986. V. 5. N 3. P. 53.
146. Химические приложения топологии и теории графов/Пер. с англ. под ред. Кинга Р. М.: Мир, 1987. 560 с.
147. Ramaswamy N.//Chem. Inf. Comput. Sci. 1987. V. 27. N 2. P. 82.
148. Ghose A. K., Crippen G. M.//J. Comput. Chem. 1986. V. 7. N 4. P. 565.
149. Young G., Lien E. J., Guo Z.//QSAR. 1986. V. 5. N 1. P. 12.
150. McFarland J. W., Gans D. J.//J. Med. Chem. 1986. V. 29. N 4. P. 505.
151. Derde M. P., Massart D. L.//Anal. Chim. acta. 1986. V. 191. P. 1.
152. Advances in Multivariate Statistical Analysis. 3/Ed. Gupta A. K. Dordrecht: Reidel, 1987. 402 р.
153. McFarland J. W., Gans D. J.//QSAR in Drug Des. Toxicol./Eds Hadzi D. et al. Amsterdam: Elsevier, 1987. P. 25.
154. Lopes de Compadre R. L., Pearlstein R. A. et al.//Ibid. P. 149.
155. Janssen L. H. M.//Ibid. P. 191.
156. Morris J. J., Hughes L. R., Taylor P. J.//Ibid. P. 204.
157. Goldblum A.//Ibid. P. 207.
158. Bruns A., Wahnsiedler W., Weitzel P. et al.//Ibid. P. 103.
159. Tichy M.//Ibid. P. 109.
160. Grunewald G. L., Sall D. J., Monn J. A.//Ibid. P. 118.
161. Peace E. A., Rord M. G., Greenwood R. et al.//Ibid. P. 331.
162. Leysen J. E., Gommeren W.//Drug dev. Res. 1986. N 8. P. 119.

163. Balaz S., Sturdic E., Augustin J.//Biophys. Chem. 1986. V. 24. P. 135.
164. Jakes S. E., Willett P.//J. Mol. Graphics. 1986. V. 4. N 1. P. 12.
165. Artificial Intelligence Applications in Chemistry./Eds Pierce T. H. et al. Washington: Amer. Chem. Soc. 1986.
166. Ayscough P. B., Chinnick S. J., Dybowksi R.//Chem. Ind. 1987. V. 15. P. 515.
167. Derde M. P., Buydens L., Guns C. et al.//Anal. Chem. 1987. V. 59. N 14. P. 1868.
168. Darvas E.//QSAR in Drug Des. Toxicol./Eds Hadzi D. et al. Amsterdam: Elsevier. 1987. P. 37.
169. Элти Дж., Кумбс М. Экспертные системы: концепции и примеры/Пер. с англ. М.: Финансы и статистика, 1987. 191 с.
170. Экспертные системы. Принципы работы и примеры/Пер. с англ. под ред. Форсайта Р. М.: Радио и связь, 1987. 224 с.
171. Построение экспертных систем/Пер. с англ. под ред. Хейеса-Рота Ф. и др. М.: Мир, 1987. 441 с.

Институт физиологически  
активных веществ АН СССР,  
Черноголовка, Московская обл.